1) Title: Independent evolution of sexual dimorphism and female-limited mimicry in swallowtail butterflies (*Papilio dardanus* and *P. phorcas*)

2) Authors: M.J.T.N. Timmermans[abe], M.J. Thompson[acf], S. Collins[d], A.P. Vogler[ab]

3) Affiliations:

a) Department of Life Sciences, Natural History Museum, London, SW7 5BD, United Kingdom

b) Department of Life Sciences, Silwood Park Campus, Imperial College London, Ascot, SL5 7PY, United Kingdom

c) Department of Zoology, Cambridge University, Downing Street, CB2 3EJ, United Kingdom

d) ABRI, P O Box 14308, Westlands 0800, Nairobi, Kenya

e) Current address: Department of Natural Sciences, Middlesex University, Hendon Campus, London, NW4 4BT, United Kingdom

f) Current address: Peterhouse, Cambridge University, Trumpington Street, CB2 1RD, United Kingdom

4) Key words: Batesian mimicry, phylogenomics, transcriptome, *Papilio phorcas*, sexual selection, polymorphism

5) Corresponding author:

Martijn J.T.N. Timmermans

m.timmermans@mdx.ac.uk

6) Running title: Polymorphism in *Papilio*

Abstract:

Several species of Swallowtail butterflies (genus *Papilio*) are Batesian mimics that express multiple mimetic female forms, while the males are monomorphic and non-mimetic. The evolution of such sex-limited mimicry may involve sexual dimorphism arising first and mimicry subsequently. Such a stepwise scenario through a non-mimetic, sexually dimorphic stage has been proposed for two closely related sexually dimorphic species; *P. phorcas,* a non-mimetic species with two female forms, and *P. dardanus*, a female-limited polymorphic mimetic species. Their close relationship indicates that female-limited polymorphism could be a shared derived character of the two species. Here we present a phylogenomic analysis of the dardanus group using 3964 nuclear loci and whole mitochondrial genomes showing that they are not sister species, and thus that the sexually-dimorphic state has arisen independently in the two species. Non-homology of the female polymorphism in both species is supported by population genetic analysis of *engrailed*, the presumed mimicry switch locus in *P. dardanus*. McDonald-Kreitman tests performed on SNPs in *engrailed* showed the signature of balancing selection in a polymorphic population of *P. dardanus*, but not in monomorphic populations, nor in the non-mimetic *P. phorcas*. Hence the wing polymorphism does not balance polymorphisms in *engrailed* in *P. phorcas*. Equally, unlike in *P. dardanus*, none of the SNPs in *P. phorcas engrailed* were associated with either female morph. We conclude that sexual dimorphism due to female polymorphism evolved independently in both species from monomorphic, non-mimetic states. While sexual selection may drive male-female dimorphism in non-mimetic species, in mimetic Papilios natural selection for protection from predators in females is an alternative route to sexual dimorphism.

Introduction:

Conspicuous morphological differences between males and females evident in many species are generally attributed to sexual selection, primarily on male traits (Kraaijeveld 2014), but the phenotypic divergence of the sexes may also be driven by natural selection. This was first advocated by A.R. Wallace based on his studies of Southeast Asian butterflies of the genus Papilio (Wallace 1865; see also Kunte 2008). Several butterfly species in this genus are prominent examples of male-female dimorphism (Kunte 2009). These species are mostly sex-limited Batesian mimics that gain protection from predation through close morphological similarity with chemically protected models while being not defended themselves (Bates 1862). Mimicry in these species is derived and often limited to the female sex, while the males are non-mimetic (Kunte 2009).

Sexual dimorphism is frequently achieved by suppressing male-selected traits in the females (Kraaijeveld 2014). In insects, this is often mediated by sex-specific splice variants of *doublesex* (*dsx*) (see e.g. Kraaijeveld 2014). This locus has been documented to determine sexual dimorphism in the two closely related Southeast Asian swallowtails *Papilio polytes* and *P. memnon* (Kunte *et al.* 2014; Nishikawa *et al.* 2015; Komata *et al.* 2016). Both species are sexually dimorphic Batesian mimics that express a mimetic colour morph in females in the presence of the dominant *dsx*(*H*) allele, whereas homozygous recessive *dsx*(*h*) females are non-mimetic and resemble the males (Kunte *et al.* 2014; Nishikawa *et al.* 2015; Komata *et al.* 2016).

The evolution of female-limited mimicry has been hypothesized to evolve through several mutational steps, starting with an initial mutation with female limited effect, for example involving the loss of male selected traits in females. After the evolution of sexual dimorphism, the female-limited phenotype might develop similarity to a toxic model species (Fisher 1927; Nicholson 1927), a process that might be followed by the evolution of additional mimetic forms in sex-limited polymorphic species, either from this initial protected state, or from the ancestral state (Turner 1984). Conversely, female-limitation of the mimetic

pattern(s) might also evolve after the evolution of mimicry, perhaps by negative frequency-dependent selection driving the loss of mimicry in males (Kunte 2009). Under the first hypothesis, sexually dimorphic traits first arise because they may be driven by sexual selection, for example exerted by female choice or male-male competition. Under the second hypothesis the dimorphic trait arises because it confers a greater advantage when limited to one sex only, i.e. as a product of natural selection. These different explanations for the causes of sex-limited inheritance of adaptive phenotypes were part of an exchange between Darwin and Wallace (Hoquet & Levandowsky 2016).

Phylogenetic reconstruction of closely related species potentially could dissect the evolutionary steps leading to sexual dimorphism and female-limited mimicry. The African Mocker Swallowtail, *P. dardanus*, and its relatives are particularly pertinent to phylogenetic studies of complex phenotypes and the evolution of polymorphic systems (Charlesworth & Charlesworth 1975; Vane-Wright *et al.* 1999; Clark *et al.* 2008; Clark & Vogler 2009). This group includes two sexually dimorphic species, *P. dardanus* and *P. phorcas*. *Papilio dardanus* exhibits numerous distinct mimetic female phenotypes throughout the species' range in sub-Saharan Africa (Thompson & Timmermans 2014), which are controlled by a single autosomal locus, termed *H*, that expresses a separate allele for each morph (Clarke & Sheppard 1959, 1960; Clarke & Sheppard 1960). Non-mimetic subspecies also exist in *P. dardanus*, specifically *P. d. meriones* and *P. d. humbloti* from Madagascar and the Comoro Islands, in which the female pattern is similar to that of the non-mimetic males ('male-like' females of Clarke & Sheppard 1960). *Papilio phorcas* has two non-mimetic female forms, one of which is similar to the monomorphic males. The two other closely related species within the group are monomorphic, including the non-mimetic *P. constantinus* and the mimetic *P. rex* (Figure 1).

Phylogenetic analyses of *P. phorcas* and *P. dardanus* potentially can elucidate the question about the evolutionary path to female-limited polymorphic mimicry (Clarke *et al.* 1991; Vane-Wright *et al.* 1999; Clark *et al.* 2008; Clark & Vogler 2009). If both species are

found to be sister species, an evolutionary route involving an initial switch to sexual dimorphism would be supported, followed by the subsequent acquisition of mimetic phenotypes in the females (Clarke *et al.* 1991; Vane-Wright *et al.* 1999). Phylogenetic trees based on mitochondrial DNA and nuclear ITS sequences have supported the conclusion that *P. phorcas* and *P. dardanus* are sister species and that *P. constantinus* is their closest relative (Vane-Wright *et al.* 1999; Clark & Vogler 2009; but see Caterino & Sperling 1999), which is consistent with the observation that the two polymorphic species form hybrids in nature (Clarke 1980; Thompson *et al.* 2011); however, this specific phylogenetic arrangement is not supported by data on the nuclear EF-1alpha locus (Vane-Wright et al. 1999), other nuclear markers (Clark & Vogler 2009), or by combinations of nuclear and mitochondrial loci (Zakharov *et al.* 2004). The common ancestry of female limited polymorphic wing patterning, and hence sexually dimorphism, was favoured by Nijhout (2003) based on presumed commonalities in wing pattern elements across the female forms in both species.

In addition to phylogenetic analyses, the increased knowledge about the molecular control of the mimetic polymorphism in *P. dardanus* permits a more direct approach to the question of common ancestry of the female limited polymorphic mimicry system. Associations between female morphs and SNPs in the *engrailed* genomic region suggest that this gene either is the wing pattern (*H*) locus of *P. dardanus*, or is tightly linked to the *H* locus (Timmermans *et al.* 2014). Morph-associated SNPs are organized into highly diverged haplotypes that differ by numerous non-synonymous substitutions, supporting the hypothesis that balancing selection maintains different alleles at this locus within polymorphic populations (Thompson *et al.* 2014; Timmermans *et al.* 2014)

A common ancestry for the female-limited polymorphism in the two species would be supported if SNPs in the orthologous *P. phorcas engrailed* region are associated with either of its two morphs, which we here test by association studies. Under such scenario, tests for adaptive evolution might detect a signal of balancing selection in *P. phorcas engrailed, as*

has been observed for a polymorphic *P. dardanus* population (Timmermans *et al.* 2014). Also, if balancing selection is the cause of the *engrailed* polymorphism in *P. dardanus*, monomorphic populations, including the male-like *P. d. meriones* race, and the Western African subspecies *P. d. dardanus* which exhibits female-limited mimicry, but is essentially monomorphic for a single mimicry morph (form *hippocoon*) should not show such evidence of selection.

We use a phylogenomic approach to re-examine the potential sister relationship of *P. dardanus* and *P. phorcas*, including the position of the monomorphic Madagascan subspecies within *P. dardanus*, and conduct a test of adaptive molecular evolution on the *engrailed* locus. The results shed light on the evolutionary progression leading to sexual dimorphism and the acquisition of multiple mimetic phenotypes limited to the females.

Material and methods:

Samples and sequence data

A phylogenomic dataset was constructed by transcriptome sequencing and by shotgun sequencing of total genomic DNA of the members of the dardanus group (*P. rex*, *P. phorcas*, *P. constantinus*, *P. dardanus dardanus*, *P. dardanus tibullus*, *P. dardanus polytrophus*, *P. dardanus cenea, P. dardanus meriones*; Figure 1). Details on starting material, read length and number of reads generated for each of the libraries are given in Supplementary table S1. Transcriptome data were generated for *P. d. cenea* (from pooled dissected wing discs from pupae), *P. d. dardanus* (thorax), *P. phorcas* (thorax), and *P. constantinus* (pupa). Total RNA was extracted using the Qiagen RNeasy mini kit (*P. d. cenea*) or the Qiagen RNA kit (the other samples). Residual DNA was removed with Ambion Turbo DNA-free (Thermofisher). Quality checks and quantification of extracted RNA was assessed using agarose gel electrophoreses, on a Qubit 2.0 Fluorometer and using a RNA chip on a Bioanalyzer (Agilent). Sequencing libraries were sequenced on the Illumina GAIIx (*P. d*

cenea; 50bp single-end sequencing) or MiSeq (other samples; 250bp paired-end sequencing) platforms.

Whole genome shotgun sequence (WGS) data were generated for *P. rex* and four additional subspecies of *P. dardanus* (Supplementary table S1; Figure 1). DNA was extracted from a small piece of thorax using the Qiagen Blood and Tissue kit. The protocol was modified slightly in that ATL buffer was replaced with a cetyltrimethylammonium bromide (CTAB) buffer (2% CTAB (weight/volume), 100 mM Tris−HCl (pH 8.0), 20 mM EDTA, 1.4 M NaCl). Prior to adding buffer AL, the CTAB solution was treated with an equal volume of chloroform:isoamyl alcohol (24:1) to remove protein. Sequencing libraries were sequenced on the Illumina GAIIx (100bp paired-end sequencing) or MiSeq (250bp or 300bp paired-end sequencing) platforms. Two further *Papilio* sequence datasets were downloaded from the GenBank SRA database (*P. glaucus* - SRR850325 and *P. polytes* - SRR850327; Zhang *et al.* 2013).

The GAIIx transcriptome reads (*P. d. cenea*) were assessed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and error-corrected using Reptile (MacManes & Eisen 2013; Yang *et al.* 2010). All other data were processed using the prinseq-lite perl script (Schmieder & Edwards 2011) using the following settings: exact_only, derep: 14, min_qual_mean: 20, ns_max_n: 0, trim_qual_right: 30.

Transcriptome and genome assembly and non-redundant reference set

The *P. d. cenea* GAIIx data was assembled using Trinity (version 2013-11-10 release; Grabherr *et al.* 2011). A range of parameters was tested for the *minimum kmer coverage* and *path reinforcement distance* parameters to obtain the highest mean and median contig length. Only contigs of more than 200 bp were retained for further analysis. Completeness of the assembly was assessed using CEGMA (Parra *et al.* 2007; Parra *et al.* 2009). A non-redundant reference sequence set was generated with CD-hit (Li & Godzik 2006) using the following settings: C= 0.95 n=0.8 r=1. This dataset was then parsed through HaMStR

(Ebersberger *et al.* 2009) to remove putative paralogous sequences using the LEP1-Cos ortholog set (Kawahara & Breinholt 2014) with *H. melpomene* as reference. HaMsTr selects best hits to a reference set and concatenates partial hits to increase total sequence length for each locus.

Phylogenomic analyses

The reference transcriptome pruned by HaMsTr was used for phylogenetic analyses. To obtain presumed homologous sequences for each of the other (sub)species, the quality controlled and de-replicated sequence reads were mapped onto the non-redundant transcriptome reference set using BWA mem (http://bio-bwa.sourceforge.net/bwa.shtml) (default settings). To allow for the more divergent reads of the outgroups to be mapped, the *P. polytes* and *P. glaucus* reads were aligned using the short read aligner BBmap (http://sourceforge.net/projects/bbmap/) (setting: local=t), which is more tolerant to sequence divergence than BWA. SAM files were converted to sorted BAM files using Samtools (Li *et al.* 2009), and Picard (v.1.117) (http://broadinstitute.github.io/picard/) was used to add read group information. GATK (https://www.broadinstitute.org/gatk/) was used to call sites different to the reference sequence (in homozygous and heterozygous states). All sites were emitted into a vcf (https://github.com/samtools/hts-specs) file, which was subsequently converted into fasta formatted sequences using a custom Perl script.

The information on the different specimen was combined into a multi-fasta alignment file. These files were subsequently combined into a single concatenated data matrix. The number of positions that were homozygous and differed from the *P. d. cenea* reference sequence was determined for each specimen and comparisons were made after correcting for missing data by dividing this number by the fraction of available data.

For phylogenetic analysis, positions with missing data for any of the biological samples were subsequently discarded and individual multi-fasta alignments with a final length <100 were removed. To obtain a suitable partitioning scheme, the alignment was analysed using

PartitionFinder (Lanfear *et al.* 2012) (starting with one partition per non-redundant transcript fragment) based on the Bayesian Information Criterion (BIC). The partitioned data matrix was subsequently used for Maximum Likelihood (ML) based phylogenetic analysis. Analyses were performed in RAxML (Stamatakis 2006) and used a GTR+Γ model for each partition. In addition, RAxML trees were generated for each of the individual alignments, again using the GTR+Γ model for each of the datasets. Branch lengths were removed from these individual phylogenetic trees ('best trees') using newick-utils (v.1.6) (Junier & Zdobnov 2010) and a "Densitree" tree plot (Bouckaert 2010) was created in R (https://www.r-project.org/) using the packages ape (Paradis *et al.* 2004) and phangorn (Schliep 2011).

In addition, full mitochondrial genomes were obtained for eight specimens of the dardanus group, either by LR-PCR (*P. d. dardanus* (Uganda), *P. phorcas*, *P. constantinus*,) following (Timmermans *et al.* 2014b), by assembly from sequenced total DNA (*P. d. dardanus* (Ghana)) as described in Gillett *et al.* (2014), or were retrieved from the above WGS datasets (*P. d. tibullus*, *P. d. polytrophus*, *P. d. meriones, P. rex*) by assembling the full dataset using Newbler (Margulies *et al.* 2005) and BLAST-based (Altschul *et al.* 1990) identification of the mitogenome sequence from the resulting contigs based on sequence similarity to an existing *P. dardanus* mitogenome sequence (JX313686; Timmermans *et al.* 2014). Finally, an incomplete mitogenome sequence for *P. d. cenea* was obtained by mapping transcriptome reads to the existing *P. dardanus* mitogenome using Geneious R8 (Kearse *et al.* 2012). The data on the nine newly generated mitogenome sequences were merged with seven existing mitogenomes (the existing *P. dardanus* sequence and six outgroup species) obtained from NCBI GenBank. Sequences of the protein coding and rRNA genes were extracted using FeatureExtract (Wernersson 2005), aligned with Transalign (Bininda-Emonds 2005) for protein coding genes and MAFFT (Katoh *et al.* 2009) for rRNA genes (E-INS-i algorithm), and combined into a single multigene data matrix. Tree searches were conducted on all positions or after masking synonymous positions in alignments of protein coding genes using the Degen script (Regier *et al.* 2010; Zwick *et al.* 2012).

PartitionFinder (Lanfear *et al.* 2012) was used to select a suitable partition scheme (starting with one partition per codon-position per gene) based on the BIC. Maximum likelihood tree searches were conducted with RAxML using a GTRCAT model for each partition.

Molecular evolution of wing-pattern candidate *engrailed*

Population-level analyses were conducted by sequencing a 424 bp portion of the *engrailed* coding region amplified with primers Pd202-Pd204 (Thompson *et al.* 2014) for *P. phorcas* (Kenya; Mount Kenya, specimen collected June 2014) and three subspecies of *P. dardanus*, including the East African *P. d. polytrophus* (Kenya; Mount Kenya, specimen collected in June 2014), the West African *P. d. dardanus* (Ghana; Aburi Botanical Gardens, specimen collected in March 2012) and the monomorphic *P. d. meriones* (Madagascar; obtained at various sites and times). Fragments were bi-directionally sequenced using Sanger technology. For each (sub)species a gene alignment was generated. Haplotypes were inferred using the "phase" algorithm in DnaSP v5 (Librado & Rozas 2009) using default settings. Inferred haplotypes were combined with data on three *P. rex* specimens (KJ458896, KJ458897 and the specimen described above) or four *P. constantinus* specimens (KJ458933, KJ458919, KJ458894, KJ458895) and the number of synonymous differences per synonymous site was calculated for each alignment using the Nei-Gojobori model (Nei & Gojobori 1986) in MEGA6 (Tamura *et al.* 2013). In order to detect signatures of selection, McDonald-Kreitman (McDonald & Kreitman 1991) tests were performed on each population sample using the MKT webserver (Egea *et al.* 2008) using default settings. For this, Jukes and Cantor (Jukes & Cantor 1969) corrections were applied. We also reanalysed data on *engrailed* from a population on Mount Kenya that was used for a previous MK test and that earlier had used *P. polytes* and *P. glaucus* for comparison (Timmermans *et al.* 2014). Timmermans et al. (2014) used a different reverse primer resulting in a 407bp fragment included here. Finally, Fisher's exact tests were performed to test whether SNPs

in *engrailed* are associated to wing phenotype in *P. phorcas*. Bonferroni corrections were applied to correct for multiple testing for these Fisher's exact tests.

Results:

Generating a phylogenomic data set

Illumina sequence data for seven specimens (Supplementary Table S1) ranged from ~5M to 50M reads. Quality filtering removed between 2% and 53% of the reads, mostly because reads were redundant, i.e. exact duplicates in particular in the RNA-seq datasets. A similar redundancy was seen in the two publicly available datasets of *P. glaucus* and *P. polytes* obtained from the SRA database, which included between 54% and 58% of redundant reads. Redundancy in the WGS datasets was significantly lower with a maximum of 12% for sample BMNH1043081 (*P. d. tibullus*). The Trinity wing disc transcriptome assembly was composed of 29184 contigs. The transcriptome was assessed for completeness against 248 core eukaryotic genes, which suggested the data is 95.5% complete and 98.7% complete if partially assembled genes were included.

CD-hit reduced the dataset to a non-redundant sequence set of 27489 contigs, which were compared to the custom ortholog set of 6568 loci of Kawahara & Breinholt (2014) using *H. melpomene* as reference. The HMM based HaMSTR pipeline identified 5408 putative homologues in the non-redundant wing transcriptome set. These loci were used for further phylogenomic analyses and homologous sequences were obtained for the seven *dardanus* group specimens and two outgroups by mapping reads onto the transcriptome reference set. Overall coverage obtained with each of the samples is given in supplementary figure S1. The gDNA based sequence sets showed a significant percentage of positions having a coverage >10x. Sequence coverage of the three RNA-seq samples was generally much lower.

Phylogenetic analyses

For each of the ten taxa (*P. d. cenea* reference + 7 *P. dardanus*-species group + 2 outgroups) sequences for the 5408 loci were merged into a concatenate of 7,387,602 bp. The number of positions that were homozygous and differed from the *P. d. cenea* reference sequence was counted for each specimen (Figure 2). As expected, the number of such sites increased with phylogenetic distance. Sites with missing data were subsequently removed and loci with a total length of 99 bp or less were also discarded, for a final set of 3964 loci (with a total length of 2,564,740 bp). Maximum Likelihood trees were obtained for each of these loci, as well as for the concatenated dataset (Figure 3). The latter analysis produced topologies with high bootstrap values and support the sister relationship of *P. constantinus* and *P. dardanus*, with *P. phorcas* as sister to both of them. It also indicated that *P. d. meriones* split at the basal node of the *P. dardanus* lineage. Phylogenetic analyses of mitogenomes revealed the same topology as obtained with the nuclear dataset, with *P. constantinus* as sister to *P. dardanus* and the more distant position of *P. phorcas* (Figure 4). Within *P. dardanus* the placement of *P. d. meriones* remained uncertain based on the mitogenomes, as it was grouped with the West African population but with low bootstrap support, while it was well separated from the polymorphic Eastern African populations.

Molecular evolution of *engrailed* and association with *P. phorcas* phenotypes

McDonald-Kreitman (MK) tests were performed on four population samples, including the highly polymorphic East African *P. d. polytrophus* (Kenya) (n=67; all females), the West African *P. d. dardanus* (Ghana) (n=13; all males), the monomorphic *P. d. meriones* (Madagascar) (n=13), and the polymorphic *P. phorcas* (Kenya) (n=15; all females, 8 of which exhibiting the male-like phenotype). In addition, data on *P. d. polytrophus* (Kenya) from an earlier study was reanalysed (Timmermans *et al.* 2014) (Table 1). The *P. d. polytrophus* (Kenya) population datasets, with multiple mimetic forms, showed a significant excess of polymorphic non-synonymous mutations in *engrailed,* consistent with long-term balancing selection (Nielsen 2005). The tests were not significant for the monomorphic *P. dardanus*

317 datasets from Ghana and Madagascar, or for *P. phorcas*, which has dimorphic, but non-
318 mimetic females. A total of 29 SNPs were observed in the *P. phorcas* alignment, and after
319 Bonferroni correction none of these were significantly associated to either of the two female
320 phenotypes.
321
322 **Discussion:**
323 Understanding the relationship of *P. dardanus* and *P. phorcas* is considered critical for
324 understanding the evolution of female-limited polymorphic mimicry (Vane-Wright *et al.* 1999).
325 Although such phylogenetic analysis does not provide direct evidence for the evolution of the
326 colour polymorphism itself, a sister relationship of these female-limited polymorphic species
327 may indicate the antiquity of this trait and thus support the evolution of female-limited
328 polymorphic mimicry in *P. dardanus* from an ancestor that was already sexually dimorphic
329 (Nijhout 2003). More distant relationships of the two species would indicate that sexual
330 dimorphism has arisen independently, possibly through the evolution of a sexual
331 monomorphic, mimetic state in *P. dardanus*, followed by the secondary loss of mimicry in the
332 males (e.g. Kunte 2009; Vane-Wright 1971). Our phylogenetic analysis does not support the
333 hypothesis of a common origin of sexual dimorphism. Separate gains are also corroborated
334 by the non-homologous mechanism of engrailed involvement, which is evident from the fact
335 that the female polymorphism is not associated to variation in *engrailed* in *P. phorcas*, unlike
336 in *P. dardanus*, and that the signature of balancing selection in *engrailed* was missing from
337 *P. phorcas*, but was clearly evident in the polymorphic populations of *P. dardanus*. These
338 results for the dardanus group are in accordance with the phylogenetic study of Kunte (2009)
339 who found no evolutionary correlation of sexual and mimetic dimorphism generally in the
340 genus *Papilio*. However, strictly speaking, we cannot exclude a scenario in which the
341 ancestor of the whole clade was sexually dimorphic, with *P. constantinus* and *P. d. meriones*
342 subsequently loosing sexual dimorphism. If so, female-limited polymorphisms would have
343 evolved secondarily, given the non-homology of the *P. phorcas* and *P. dardanus*

phenotypes, but this scenario is entirely hypothetical and not supported by any specific evidence.

The independent origins of the female limited mimetic polymorphisms in various Papilios (Kunte 2009) might suggest a different underlying genetic basis in each mimetic species. The recent discovery that *dsx* acts as the switch locus in two Southeast Asian swallowtails *P. polytes* (Kunte *et al.* 2014; Nishikawa *et al.* 2015) and *P. memnon* (Komata *et al.* 2016) is intriguing, especially as sequence analysis suggests that the mimetic polymorphism evolved independently in both species. Yet, the lack of involvement of *engrailed* in those species, supports the non-homology of the polymorphic mimicry systems across the wider genus *Papilio*. For *P. phorcas*, the morph-determining locus remains unknown, but all evidence obtained here argues against the involvement of *engrailed*.

Any hypothesis for the evolution of mimetic forms in *P. dardanus* has to take into account that the evolution of male and female patterns is uncoupled, and potentially involves different mechanism; *engrailed* determines the female morph, but it may not determine the sexual dimorphism *per se*, while *vice versa* we do not know what role *engrailed* plays in determining the male pattern, and if it has a role at all. The male patterns in both *P. dardanus* and *P. phorcas* are curious because they are very different from each other and also differ greatly from the presumed ancestral yellow-banded pattern in *P. constantinus* and other related species (Figure 1). This suggests a shift to derived, but non-mimetic male phenotypes in each case, possibly driven by sexual selection, as pointed out by Nijhout (2003).

Molecular analyses may inform in greater detail about the separate determination of female morph and male-female dimorphism. In the single case of *P. polytes* studied to date, maleness appears to be achieved by suppressing the female phenotype. Specifically, RNAi knockdown experiments that silenced the dominant *dsx* allele in *P. polytes* showed that this allele not only specifies the mimetic female pattern but at the same time suppresses the non-mimetic (male-like) phenotype (Nishikawa *et al.* 2015), supporting a dual function of the

mimicry *H* alleles in specifying the female morph and the sexual dimorphism. In analogy, the specification of the female morph in *P. dardanus* is through *engrailed.* However, the sexual dimorphism might be determined by a different locus, whose identity we don't know, but based on its role in other insects (Kraaijeveld 2014) could quite possibly be the *dsx* locus.

This hypothesis about what determines male patterns also bears on the question about the *H* status of the male-like females in *P. dardanus* and their evolution. If the male-like females result simply from suppression of maleness, a mutational loss-of-function in the male-suppressing locus would produce a male-like phenotype in the females, i.e. male-like females could have evolved secondarily. This possibility is unlikely for the monomorphic Indian Ocean subspecies *P. d. meriones* that apparently express the plesiomorphic state (sexual monomorphism) of the *P. dardanus* lineage. A male-like female phenotype, however, is also found in several polymorphic African mainland populations (Clarke & Sheppard 1960; Thompson & Timmermans 2014), which may have originated from reversals, by losing the repression of the mimicry male phenotype due to a mutation in the male-supressing locus. Importantly, in both cases the male pattern is not a springboard from which female forms diversify.

Taken together, current evidence reinforces the general notion that the mimetic patterns are derived and the male pattern is ancestral. Initially Clarke *et al.* (1985) inferred the ancestry of the male patterns based on wing pattern dominance hierarchies. As expected for an ancestral phenotype, crosses performed by these authors revealed that the male pattern (as displayed by male-like females) was recessive to the other female forms in most cases (Clarke *et al.* 1985). There may be problems with the interpretation of these dominance patterns, in particular involving crosses between the monomorphic Malagasy and polymorphic African mainland individuals, because the phenotypes in the offspring are not expressed cleanly. However, together with the increasingly better understanding of the underlying genetic mechanism, including RNAi experiments in *P. polytes* (Nishikawa *et al.* 2015) showing the suppression of maleness via the presumed derived *dsx* alleles, there is

sufficient evidence for the non-mimetic state to be plesiomorphic, which is retained in the males, and from which the mimicry pattern is derived only in the females driven by natural selection for the avoidance of predation. This does not mean, however, that the male pattern cannot undergo changes also, as is clearly the case in *P. dardanus* and *P. phorcas* whose males do not resemble the presumed ancestral form (Vane-Wright et al. 1991).

The alternative possibility for the evolution of sexual dimorphism driven by sexual selection on the male phenotype seems to be refuted for *P. dardanus*, and also seems less common based on the phylogenetic reconstructions of male and female-limited traits in other (non-mimetic) butterflies (Oliver & Monteiro 2011). However, the case of *P. phorcas*, which differs in particular by the absence of mimicry, may yet be explained by the divergence of males, possibly driven by sexual selection, from which females are secondarily derived with similar male-like phenotypes.

Acknowledgements:

The authors would like to thank Ashanti African Tours, S. Sáfián (Butterfly Conservation Ghana), L. Kirkpatrick and O. Brattström for help with sampling of butterflies and/or for providing butterfly tissue and R. Massang for very kindly granting permission to collect *P. dardanus* in Aburi Botanical Gardens (Republic of Ghana). The authors also thank C. Jiggins for constructive discussions on transcriptome assembly and phylogenomic analysis and K. Kraaijeveld, J. Paps and two anonymous reviewers for their valuable comments on our manuscript. The NHM London and NBAF Edinburgh generated sequencing libraries and performed the Illumina sequencing. This study was funded by NERC Postdoctoral Fellowship NE/I021578/1 (to MJTNT), NERC NE/F006225/1 (to APV), and a NERC doctoral fellowship to MJT.

Data Accessibility:

All Illumina sequence datasets are stored in the Short Read Archive (http://www.ncbi.nlm.nih.gov/sra) under accession numbers SAMN05819004-SAMN05819010. Sanger sequences and newly generated mitochondrial genomes have been deposited in NCBI GenBank under accession numbers KX034410-KX034516 and KX033351-KX033358.

Author Contributions:

Designed the study: MJTNT, MJT, APV. Collected material: SC. Analysed the data: MJTNT, MJT. Wrote manuscript: MJTNT and APV. All authors read and approved the final version of the manuscript.

Figure legends:

Figure 1: A) Species of the *dardanus* group in the current study, including the two polymorphic species *P. phorcas* and *P. dardanus*. The male of *P. dardanus* is shown in the bottom left corner. Male(-like) pattern of *P. phorcas* are shown on the left. B) The geographical distribution of the five subspecies of *P. dardanus* that were analysed.

Figure 2: The number of positions (x1000) that differed from the *P. d. cenea* reference sequence and were homozygous. Uncorrected: direct count, a total of 7,387,602 sites were investigated, positions with missing data were ignored. Corrected: Number corrected by dividing it by the fraction of missing data. A) Intraspecific comparisons, B) interspecific dardanus group comparisons, C) comparisons with outgroups.

Figure 3: A) Densitree graph visualising 3964 Maximum Likelihood phylogenetic trees (one for each transcript fragment). B) Maximum Likelihood tree based on concatenated dataset with no missing data (2,564,740 bp). Bootstrap values are shown on nodes. Scale bar indicates expected changes per site.

Figure 4: Phylogenetic relationships within the *dardanus* group based on full mitochondrial genomes. Bootstrap values are shown on nodes. Scale bar indicates expected changes per site.

Tables:

Table 1: Distances and McDonald-Kreitman test for the engrailed locus. Results from analyses with *P. rex* and *P. constantinus* are given for each (sub-)species. Divergence is defined as the number of (non-)synonymous differences per (non-)synonymous site. For the McDonald-Kreitman tests (MK test) synonymous and non-synonymous differences (fixed and variable) after Jukes and Canter (Jukes & Cantor 1969) correction are given. N) Number of specimen. (*) Data from Timmermans et al. 2014). (**) Female form *P. dardanus* f. *dionysos* occurs at low frequency in this population

Supplementary data:

Supplementary Table S1: Detailed information on sequence datasets, including read length, number of sequences and Quality Control. Paired reads were generated and the first column in each field refers to the forward reads (R1) and the second column to the reverse reads (R2). Dataset SRR850325 and SRR850327 from Zhang *et al.* (2013).

Supplementary Figure S1: Percentage of bases in the transcriptome dataset with the coverage given on the x-axis. For each of the nine sequence sets the reads were mapped onto the non-redundant wing transcriptome dataset and coverage information was obtained using Samtools “depth” function.

References:

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of molecular biology*, **215**, 403–410.

Bates HW (1862) Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. *Transactions of the Linnean Society London*, **23**, 495–566.

Bininda-Emonds ORP (2005) transAlign: using amino acids to facilitate the multiple alignment of protein-coding DNA sequences. *BMC bioinformatics*, **6**.

Bouckaert RR (2010) DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics*, **26**, 1372–1373.

Caterino MS, Sperling FAH (1999) Papilio phylogeny based on mitochondrial cytochrome oxidase I and II genes. *Mol Phylogenet Evol*, **11**.

Charlesworth D, Charlesworth B (1975) Theoretical genetics of Batesian mimicry. 2. Evolution of supergenes. *Journal of Theoretical Biology*, **55**, 305–324.

Clarke CA (1980) *Papilio nandina*, a probable hybrid between *Papilio dardanus* and *Papilio phorcas*. *Systematic Entomology*, **5**, 49–57.

Clarke CA, F. M. M. Clarke, S. C. Collins, A. C. L. Gill, J. R. G. Turner (1985) Male-like females, mimicry and transvestism in swallowtail butterflies. *Systematic Entomology*, **10**, 257–283.

Clarke SC, I. J. Gordon, R. I. Vane-Wright, C. R. Smith (1991) Phylogenetic relationships of three African swallowtail butterflies, *Papilio dardanus, P. phorcas* and *P. constantinus*: new data from hybrids (Lepidoptera: Papilionidae). *Systematic Entomology*, **16**, 257–273.

Clarke CA, P. M. Sheppard (1959) The genetics of *Papilio dardanus* Brown. I. Race *cenea* from South Africa. *Genetics*, **44**, 1347–1358.

Clarke CA, P. M. Sheppard (1960) The genetics of *Papilio dardanus* Brown. II. Races *dardanus*, *polytrophus*, *meseres*, and *tibullus*. *Genetics*, **45**, 439–456.

Clarke CA, Sheppard PM (1960) The genetics of *Papilio dardanus* Brown. III. Race *antinorii* from Abyssinia and race *meriones* from Madagascar. *Genetics*, **45**, 683–698.

508 Clark R, S. M. Brown, S. C. Collins *et al.* (2008) Colour pattern specification in the Mocker Swallowtail
509 *Papilio dardanus*: the transcription factor *invected* is a candidate for the mimicry locus *H*.
510 *Proceedings of the Royal Society B*, **275**, 1181–1188.
511 Clark R, Vogler AP (2009) A phylogenetic framework for wing pattern evolution in the mimetic
512 Mocker Swallowtail Papilio dardanus. *Molecular Ecology*, **18**, 3872–3884.
513 Ebersberger I, Strauss S, Haeseler A (2009) HaMStR: Profile hidden markov model based search for
514 orthologs in ESTs. *BMC Evolutionary Biology*, **9**, 1–9.
515 Egea R, Casillas S, Barbadilla A (2008) Standard and generalized McDonald–Kreitman test: a website
516 to detect selection by comparing different classes of DNA sites. *Nucleic Acids Research*, **36**,
517 W157–W162.
518 Fisher RA (1927) On some objections to mimicry theory; statistical and genetic. *Transactions of the*
519 *Royal Entomological Society*, **75**, 269–274.
520 Ford EB (1936) The genetics of *Papilio dardanus* Brown (Lep.). *Transactions of the Royal*
521 *Entomological Society London*, **85**, 435–466.
522 Gillett CPDT, Crampton-Platt A, Timmermans MJTN *et al.* (2014) Bulk de novo mitogenome assembly
523 from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea).
524 *Molecular biology and evolution*, **31**, 2223–2237.
525 Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq
526 data without a reference genome. *Nat Biotech*, **29**, 644–652.
527 Hoquet T, Levandowsky M (2016) Utility vs Beauty: Darwin, Wallace and the Subsequent History of
528 the Debate on Sexual Selection. In: *Current Perspectives on Sexual Selection, History,*
529 *Philosophy and Theory of the Life Sciences*, pp. 19–28. Springer Science+Business Media,
530 Dordrecht.
531 Jukes TH, Cantor CR (1969) Evolution of protein molecules. In: *Mammalian Protein Metabolism* (ed
532 Munro H. N). Academic Press, New York.

533 Junier T, Zdobnov EM (2010) The Newick utilities: high-throughput phylogenetic tree processing in
534 the Unix shell. *Bioinformatics*, **26**, 1669–1670.
535 Katoh K, G. Asimenos, H. Toh (2009) Multiple Alignment of DNA Sequences with MAFFT. In: *Methods*
536 *in Molecular Biology. Volume 537* (ed Posada D), pp. 39–64. Humana Press.
537 Kawahara AY, Breinholt JW (2014) Phylogenomics provides strong evidence for relationships of
538 butterflies and moths. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**.
539 Kearse M, Moir R, Wilson A *et al.* (2012) Geneious Basic: an integrated and extendable desktop
540 software platform for the organization and analysis of sequence data. *Bioinformatics*
541 *(Oxford, England)*, **28**, 1647–1649.
542 Komata S, Lin C-P, Iijima T, Fujiwara H, Sota T (2016) Identification of doublesex alleles associated
543 with the female-limited Batesian mimicry polymorphism in Papilio memnon. *Scientific*
544 *Reports*, **6**, 34782.
545 Kraaijeveld K (2014) Reversible Trait Loss: The Genetic Architecture of Female Ornaments. *Annual*
546 *Review of Ecology, Evolution, and Systematics*, **45**, 159–177.
547 Kunte K (2008) Mimetic butterflies support Wallace's model of sexual dimorphism. *Proceedings of*
548 *the Royal Society B-Biological Sciences*, **275**, 1617–1624.
549 Kunte K (2009) The diversity and evolution of Batesian mimicry in Papilio swallowtail butterflies.
550 *Evolution*, **63**, 2707–2716.
551 Kunte K, Zhang W, Tenger-Trolander A *et al.* (2014) doublesex is a mimicry supergene. *Nature*,
552 **507,** 229–232
553 Lanfear R, Calcott B, Ho SYW, Guindon S (2012) Partitionfinder: combined selection of partitioning
554 schemes and substitution models for phylogenetic analyses. *Molecular biology and*
555 *evolution*, **29**, 1695–1701.
556 Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism
557 data. *Bioinformatics*, **25**.

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, **22**, 1658–9.

Li H, Handsaker B, Wysoker A *et al.* (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, **25**, 2078–2079.

MacManes MD, Eisen MB (2013) Improving transcriptome assembly through error correction of high-throughput sequence reads. *PeerJ*, e113.

Margulies M, Egholm M, Altman WE *et al.* (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, **437**, 376–380.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, **351**, 652–654.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, **3**, 418–426.

Nicholson AJ (1927) A new theory of mimicry in insects. *Australian Journal of Zoology*, **5**, 10–104.

Nielsen R (2005) Molecular signatures of natural selection. *Annual Review Of Genetics*, **39**, 197–218.

Nijhout HF (2003) Polymorphic mimicry in *Papilio dardanus*: mosaic dominance, big effects, and origins. *Evolution and Development*, **5**, 579–592.

Nishikawa H, Iijima T, Kajitani R *et al.* (2015) A genetic mechanism for female-limited Batesian mimicry in Papilio butterfly. *Nat Genet*, **47**, 405–409.

Oliver JC, Monteiro A (2011) On the origins of sexual dimorphism in butterflies. *Proceedings of the Royal Society of London B: Biological Sciences*, **278**, 1981–1988.

Paradis E, Claude J, Strimmer K (2004) APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*, **20**, 289–290.

Parra G, Bradnam K, Korf I (2007) CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics*, **23**, 1061–1067.

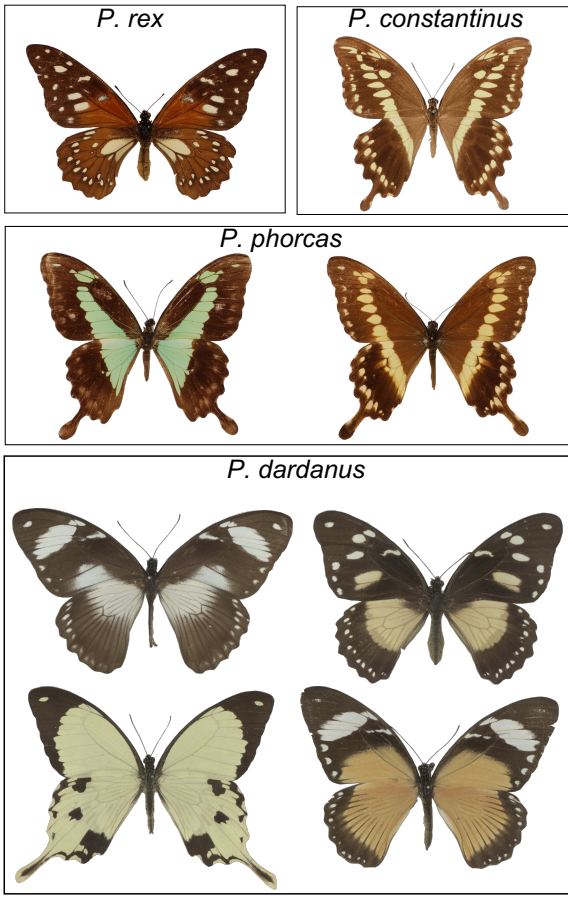
582 Parra G, Bradnam K, Ning Z, Keane T, Korf I (2009) Assessing the gene space in draft genomes.
583 *Nucleic Acids Research*, **37**, 289–297.
584 Regier JC, Shultz JW, Zwick A *et al.* (2010) Arthropod relationships revealed by phylogenomic analysis
585 of nuclear protein-coding sequences. *Nature*, **463**, 1079–U98.
586 Schliep KP (2011) phangorn: phylogenetic analysis in R. *Bioinformatics*, **27**, 592–593.
587 Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets.
588 *Bioinformatics*, **27**, 863–864.
589 Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with
590 thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.
591 Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics
592 Analysis version 6.0. *Molecular biology and evolution*, **30**, 2725–2729.
593 Thompson MJ, Timmermans MJTN (2014) Characterising the phenotypic diversity of Papilio
594 dardanus wing patterns using an extensive museum collection. *PLoS ONE*, **9**, e96815.
595 Thompson MJ, Timmermans MJ, Jiggins CD, Vogler AP (2014) The evolutionary genetics of highly
596 divergent alleles of the mimicry locus in Papilio dardanus. *BMC Evolutionary Biology*, **14**, 1–
597 13.
598 Thompson MJ, Vane-Wright RI, Timmermans MJTN (2011) HYBRID ORIGINS: DNA TECHNIQUES
599 CONFIRM THAT PAPILIO NANDINA IS A SPECIES HYBRID (PAPILIONIDAE). *Journal of the*
600 *Lepidopterists Society*, **65**, 199–201.
601 Timmermans MTJN, Baxter SW, Clark R *et al.* (accepted) Comparative genomics of the mimicry
602 switch in Papilio dardanus. *Proceedings of the Royal Society B-Biological Sciences*.
603 Timmermans MTJN, Baxter SW, Clark R *et al.* (2014a) Comparative genomics of the mimicry switch in
604 Papilio dardanus. *Proceedings of the Royal Society B-Biological Sciences*, **281**.
605 Timmermans MJTN, Lees DC, Simonsen TJ (2014b) Towards a mitogenomic phylogeny of
606 Lepidoptera. *Molecular phylogenetics and evolution*, **79**, 169–178.

607 Turner JRG (1984) Mimicry: the palatability spectrum and its consequences. In: *The biology of*
608 *butterflies* (eds Vane-Wright RI, P. R. Ackery), pp. 141–161. Academic Press, New York.
609 Vane-Wright RI (1971) The systematics of Drusillopsis Oberthür (Satyrinae) and the supposed
610 amathusiid Bigaena van Eecke (Lepidoptera: Nymphalidae), with some observations on
611 Batesian mimicry. *Transactions of the Royal Entomological Society of London*, **123**, 97–123.
612 Vane-Wright RI, D. C. Raheem, A. Cieslak, A. P. Vogler (1999) Evolution of the mimetic African
613 swallowtail butterfly *Papilio dardanus*: molecular data confirm relationships with *P. phorcas*
614 and *P. constantius*. *Biological Journal of the Linnean Society*, **66**, 215–229.
615 Wallace AR (1865) I. On the Phenomena of Variation and Geographical Distribution as illustrated by
616 the Papilionidæ of the Malayan Region. *Transactions of the Linnean Society of London*, **25**, 1–
617 71.
618 Wernersson R (2005) FeatureExtract--extraction of sequence annotation made easy. *Nucleic acids*
619 *research*, **33**, W567–569.
620 Yang X, Dorman KS, Aluru S (2010) Reptile: representative tiling for short read error correction.
621 *Bioinformatics*, **26**, 2526–2533.
622 Zakharov EV, Caterino MS, Sperling FAH (2004) Molecular phylogeny, historical biogeography, and
623 divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera :
624 Papilionidae). *Systematic Biology*, **53**, 193–215.
625 Zhang W, Kunte K, Kronforst MR (2013) Genome-Wide Characterization of Adaptation and
626 Speciation in Tiger Swallowtail Butterflies Using De Novo Transcriptome Assemblies.
627 *Genome Biology and Evolution*, **5**, 1233–1245.
628 Zwick A, Regier JC, Zwickl DJ (2012) Resolving Discrepancy between Nucleotides and Amino Acids in
629 Deep-Level Arthropod Phylogenomics: Differentiating Serine Codons in 21-Amino-Acid
630 Models. *PLoS ONE*, **7**.
631

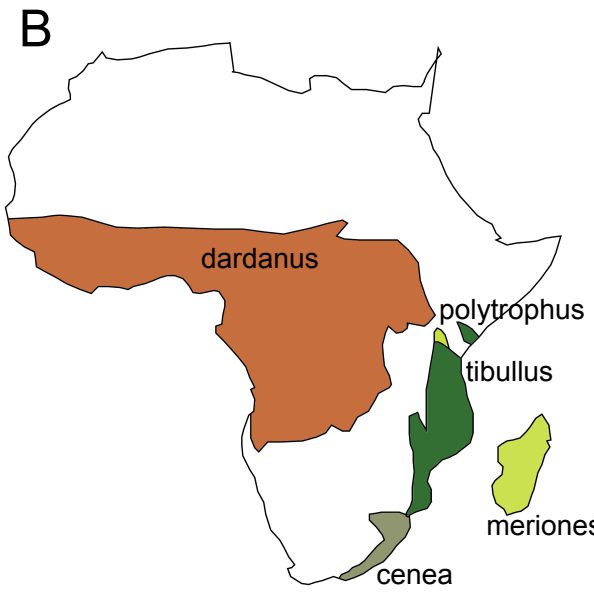
(Sub)species	Country	Female phenotype	N	to *P. rex* (r) or *P. constantinus* (c)	Non-synonymous divergence	Net non-synonymous divergence	Synonymous divergence	Net synonymous divergence	MK test				
									Synonymous Divergence	Synonymous Polymorphism	Non-synonymous Divergence	Non-synonymous Polymorphism	Fisher's exact test P-value
*P. dardanus polytrophus**	Kenya	Polymorphic - mimetic	73	r	0.025	0.016	0.278	0.234	20.89	36	3.01	36	**0.001**
				c	0.030	0.019	0.246	0.191	14.42	42	4.03	41	**0.031**
P. dardanus polytrophus	Kenya	Polymorphic - mimetic	67	r	0.022	0.016	0.272	0.236	23.34	31	3.01	27	**0.001**
				c	0.028	0.019	0.243	0.195	16.78	37	4.03	33	**0.023**
P. phorcas	Kenya	Polymorphic.not mimetic	15	r	0.032	0.027	0.328	0.273	23.34	31	8.13	10	0.887
				c	0.032	0.025	0.267	0.200	16.78	37	7.1	16	0.967
*P. dardanus dardanus***	Ghana	Monomorphic - mimetic	13	r	0.020	0.015	0.254	0.226	21.97	18	4.03	11	0.062
				c	0.025	0.018	0.232	0.192	19.34	24	5.05	17	0.086
P. dardanus meriones	Madagascar	Monomorphic - not mimetic	13	r	0.021	0.019	0.282	0.251	26.14	25	6.07	5	0.821
				c	0.026	0.021	0.258	0.215	19.34	31	6.07	11	0.834

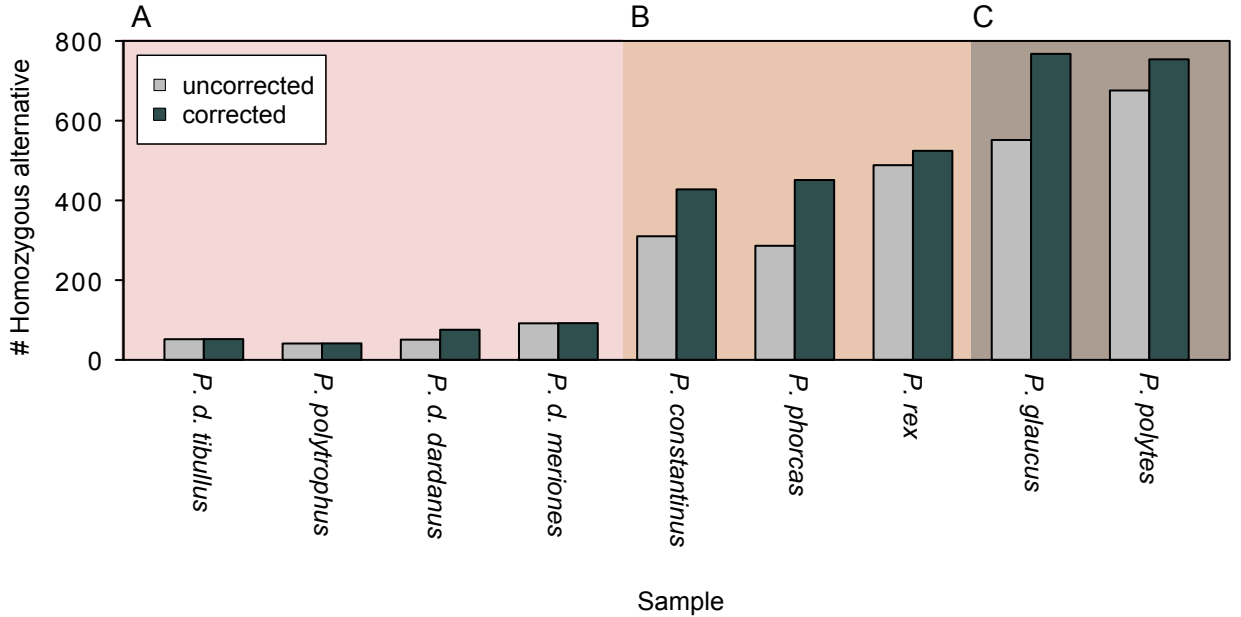
Table 1: Distances and McDonald-Kreitman test for the *engrailed* locus. Results from analyses with *P. rex* and *P. constantinus* are given for each (sub-)species. Divergence is defined as the number of (non-)synonymous differences per (non-)synonymous site. For the McDonald-Kreitman tests (MK test) synonymous and non-synonymous differences (fixed and variable) after Jukes and Canter (Jukes & Cantor 1969) correction are given. N) Number of specimen. (*) Data from Timmermans *et al.* 2014). (**) Female form *P. dardanus f. dionysos* occurs at low frequency in this population.

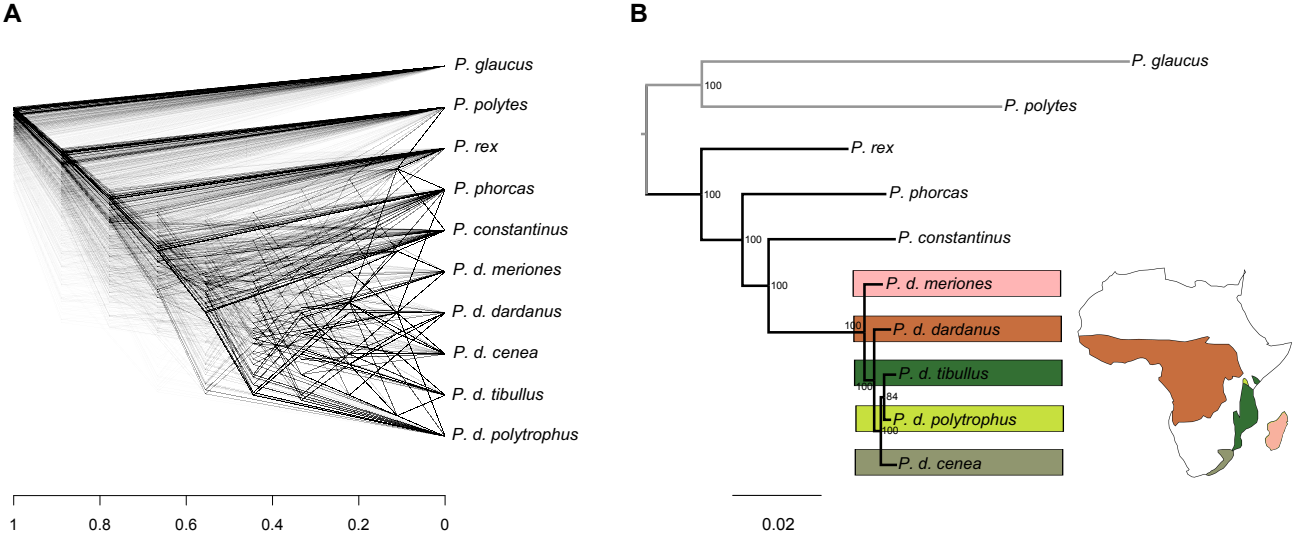
A

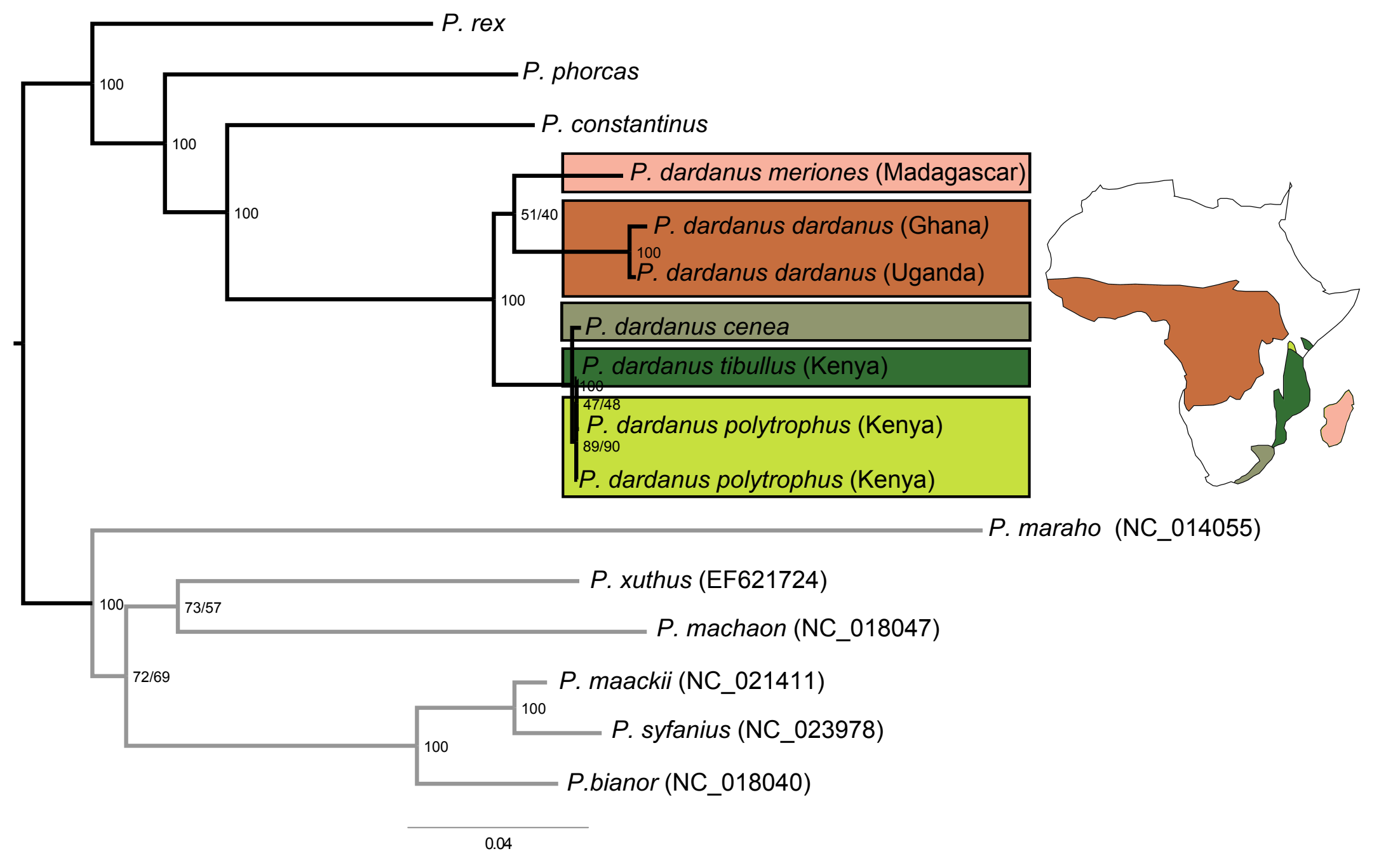

P. rex
P. constantinus
P. phorcas
P. dardanus

B


dardanus
polytrophus
tibullus
meriones
cenea


A
B
C
800
600
400
200
0
Homozygous alternative
uncorrected
corrected
P. d. tibullus
P. polytrophus
P. d. dardanus
P. d. meriones
P. constantinus
P. phorcas
P. rex
P. glaucus
P. polytes
Sample


A
P. glaucus
P. polytes
P. rex
P. phorcas
P. constantinus
P. d. meriones
P. d. dardanus
P. d. cenea
P. d. tibullus
P. d. polytrophus
1
0.8
0.6
0.4
0.2
0
B
P. glaucus
P. polytes
P. rex
P. phorcas
P. constantinus
P. d. meriones
P. d. dardanus
P. d. tibullus
P. d. polytrophus
P. d. cenea
100
100
100
100
100
100
84
100
0.02


P. rex
P. phorcas
P. constantinus
P. dardanus meriones (Madagascar)
P. dardanus dardanus (Ghana)
P. dardanus dardanus (Uganda)
P. dardanus cenea
P. dardanus tibullus (Kenya)
P. dardanus polytrophus (Kenya)
P. dardanus polytrophus (Kenya)
P. maraho (NC_014055)
P. xuthus (EF621724)
P. machaon (NC_018047)
P. maackii (NC_021411)
P. syfanius (NC_023978)
P.bianor (NC_018040)
100
100
100
51/40
100
100
100
47/48
89/90
100
73/57
72/69
100
100
0.04